A Ruelle Operator for continuous time Markov Chains

Alexandre Baraviera * Ruy Exel † Artur O. Lopes‡

March 17, 2008

Abstract

We consider a generalization of the Ruelle theorem for the case of continuous time problems. We present a result which we believe is important for future use in problems in Mathematical Physics related to C^*-Algebras.

We consider a finite state set S and a stationary continuous time Markov Chain X_t, $t \geq 0$, taking values on S. We denote by Ω the set of paths w taking values on S (the elements w are locally constant with left and right limits and are also right continuous on t). We consider an infinitesimal generator L and a stationary vector p_0. We denote by P the associated probability on $(\Omega, \mathcal{B})$. All functions f we consider bellow are in the set $\mathcal{L}^\infty(P)$.

From the probability P we define a Ruelle operator $\mathcal{L}^t, t \geq 0$, acting on functions $f : \Omega \to \mathbb{R}$ of $\mathcal{L}^\infty(P)$. Given $V : \Omega \to \mathbb{R}$, such that is constant in sets of the form $\{X_0 = c\}$, we define a modified Ruelle operator $\tilde{\mathcal{L}}^t_V, t \geq 0$, in the following way: there exist a certain f_V such that for each t we consider the operator acting on g given by

$$\tilde{\mathcal{L}}^t_V(g)(w) = [\,\frac{1}{f_V}\mathcal{L}^t(e^{\int_0^t (V\circ\Theta_s)(.)ds}\, g\, f_V)\,]\,(w)$$

We are able to show the existence of an eigenfunction u and an eigen-probability ν_V on Ω associated to $\tilde{\mathcal{L}}^t_V, t \geq 0$.

We also show the following property for the probability ν_V: for any integrable $g \in \mathcal{L}^\infty(P)$ and any real and positive t

$$\int e^{-\int_0^t (V\circ\Theta_s)(.)ds}\;[\,(\tilde{\mathcal{L}}^t_V\,(g)\,)\,\circ\theta_t\,]d\nu_V = \int g d\nu_V$$

This equation generalize, for the continuous time Markov Chain, a similar one for discrete time systems (and which is quite important for understanding the KMS states of certain C^*-algebras).

*Instituto de Matemática, UFRGS, Porto Alegre 91501-970, Brasil. Beneficiary of grant given by CAPES - PROCAD.

†Departamento de Matemática, UFSC, Florianopolis 88040-900, Brasil. Partially supported by CNPq, Instituto do Milenio, PRONEX - Sistemas Dinâmicos.

‡Instituto de Matemática, UFRGS, Porto Alegre 91501-970, Brasil. Partially supported by CNPq, Instituto do Milenio, PRONEX - Sistemas Dinâmicos, Beneficiary of grant given by CAPES - PROCAD. fax number (55)51-33087301

1 Introduction

We want to extend the concept of Ruelle operator to continuous time Markov Chains. In order to do that we need a probability a priori on paths. This fact is not explicit in the discrete time case (thermodynamic formalism) but it is necessary here.

We consider a continuous time stochastic process: the sample paths are functions of the positive real line $\mathbb{R}_+ = \{t \in \mathbb{R} \colon t \geq 0\}$ taking values in a finite set S with n elements, that we denote by $S = \{1, 2, \ldots, n\}$. Now, consider a n by n real matrix L such that:

1) $0 < -L_{ii}$, for all $i \in S$,

2) $L_{ij} \geq 0$, for all $i \neq j$, $i \in S$,

3) $\sum_{i=1}^{n} L_{ij} = 0$ for all fixed $j \in S$.

We point out that, by convention, we are considering column stochastic matrices and not line stochastic matrices (see [N] section 2 and 3 for general references).

We denote by $P^t = e^{tL}$ the semigroup generated by L. The left action of the semigroup can be identified with an action over functions from S to $\mathbb{R}$ (vectors in $\mathbb{R}^n$) and the right action can be identified with action on measures on S (also vectors in $\mathbb{R}^n$).

The matrix e^{tL} is column stochastic, since from the assumptions on L it follows that

$$(1, \ldots, 1)e^{tL} = (1, \ldots, 1)(I + tL + \frac{1}{2}t^2L^2 + \cdots) = (1, \ldots, 1).$$

It is well known that there exist a vector of probability $p_0 = (p_0^1, p_0^2, ..., p_0^n) \in \mathbb{R}^n$ such that $e^{tL}(p_0) = P^t p_0 = p^0$ for all $t > 0$. The vector p_0 is a right eigenvector of e^{tL}. All entries p_0^i are *strictly positive*, as a consequence of hypothesis 1.

Now, let's consider the space $\tilde{\Omega} = \{1, 2, ..., n\}^{\mathbb{R}_+}$ of all functions from $\mathbb{R}_+$ to S. In principle this seems to be enough for our purposes, but technical details in the construction of probability measures on such a space force us to use a restriction: we consider the space $\Omega \subset \tilde{\Omega}$ as the set of right-continuous functions from $\mathbb{R}_+$ to S, which also have left limits in every $t > 0$. These functions are constant in intervals (closed in the left and open in the right). In this set we consider the sigma algebra $\mathcal{B}$ generated by the cylinders of the form

$$\{w_0 = a_0, w_{t_1} = a_1, w_{t_2} = a_2, ..., w_{t_r} = a_r\},$$

where $t_i \in \mathbb{R}_+$, $r \in \mathbb{Z}^+, a_i \in S$ and $0 < t_1 < t_2 < \ldots < t_r$. It is possible to endow Ω with a metric, the Skorohod-Stone metric d, which makes Ω complete and separable ([EK] section 3.5), but the space is not compact.

Now we can introduce a continuous time version of the shift map as follows: we define for each fixed $s \in \mathbb{R}_+$ the $\mathcal{B}$-measurable transformation $\Theta_s : \Omega \to \Omega$ given by $\Theta_s(w_t) = w_{t+s}$ (we remark that Θ_s is also a continuous transformation with respect to the Skorohod-Stone metric d).

For L and p_0 fixed as above we denote by P the probability on the sigma-algebra $\mathcal{B}$ defined for cylinders by

$$P(\{w_0 = a_0, w_{t_1} = a_1, ..., w_{t_r} = a_r\}) = P^{t_r - t_{r-1}}_{a_r a_{r-1}} ... P^{t_2 - t_1}_{a_2 a_1} P^{t_1}_{a_1 a_0} p_0^{a_0}.$$

For further details of the construction of this measure we refer the reader to [B].

The probability P on $(\Omega, \mathcal{B})$ is stationary in the sense that for any integrable function f, and, any $s \geq 0$

$$\int f(w)dP(w) = \int (f \circ \Theta_s)dP(w).$$

From now on the Stationary Process defined by P is denoted by X_t and all functions f we consider are in the set $\mathcal{L}^\infty(P)$.

There exist a version of P such that for a set of full measure we have that all sample elements w are locally constant on t, with left and right limits, and w is right continuous on t. We consider from now on such probability P acting on this space.

From P we are able to define a continuous time Ruelle operator $\mathcal{L}^t$, $t > 0$, acting on functions $f : \Omega \to \mathbb{R}$ of $\mathcal{L}^\infty(P)$. It's also possible to introduce the endomorphism $\alpha_t \colon \mathcal{L}^\infty(P) \to \mathcal{L}^\infty(P)$ defined as

$$\alpha_t(\varphi) = \varphi \circ \Theta_t \ , \qquad \forall \varphi \in \mathcal{L}^\infty(P).$$

We relate in the next section the conditional expectation with respect to the σ-algebras $\mathcal{F}_t^+$ with the operators $\mathcal{L}^t$ and α_t, as follows:

$$[\mathcal{L}^t(f)](\Theta_t) = E(f|\mathcal{F}_t^+).$$

Given $V : \Omega \to \mathbb{R}$, such that it is constant in sets of the form $\{X_0 = c\}$ (i.e., V depends only on the value of $w(0)$), we consider a Ruelle operator family $\tilde{\mathcal{L}}_V^t$, for all $t > 0$, given by

$$\tilde{\mathcal{L}}_V^t(g)(w) = [\, \frac{1}{f_V} \mathcal{L}^t (e^{\int_0^t (V \circ \Theta_s)(.) ds} \, g \, f_V) \,] \, (w),$$

for any given g, where f_V is a fixed function.

We are able to show the existence of an eigen-probability ν_V on Ω, for the family $\tilde{\mathcal{L}}_V^t$, for all $t > 0$, such that satisfies:

Theorem A. *For any integrable $g \in \mathcal{L}^\infty(P)$ and any positive t*

$$\int [\ \frac{1}{e^{\int_0^t (V \circ \Theta_s)(.) ds} \ f_V} \ ] \ [\ \mathcal{L}^t \ (\ [\ e^{\int_0^t (V \circ \Theta_s)(.) ds} f_V] \ \ g \ ) \circ \Theta_t \ ] \ d\nu_V = \int g \, d\nu_V.$$

The above functional equation is a natural generalization (for continuous time) of the similar one presented in Theorem 7.4 in [EL1] and proposition 2.1 in [EL2] .

In [EL1] and [EL2] the important probability in the Bernoulli space is an eigen-probability ν for the Ruelle operator associated to a certain potential $V = \log H$: $\{1, 2, .., d\}^{\mathbb{N}} \to \mathbb{R}$. This probability ν satisfies: for any $m \in \mathbb{N}$, $g \in C(X)$,

$$\int \lambda_m \, E_m(\lambda_m^{-1} \, g) d\nu = \int g \, d\nu,$$

where

$$\lambda_m(x) = \frac{H^{\beta \, [m]}}{\Lambda^{[m]}}(x) = \frac{H^\beta(x) \, H^\beta(\sigma(x)) \, ... \, H^\beta(\sigma^{n-1}(x))}{\Lambda^{[m]}(x)} =$$

$$\frac{e^{\log(H^\beta(x)) + \log(H^\beta(\sigma(x))) + \ldots + \log(H^\beta)(\sigma^{n-1}(x))}}{\Lambda^{[m]}(x)},$$

and σ is the shift on the Bernoulli space $\{1, 2, .., d\}^{\mathbb{N}}$. Here $E_m(f) = E(f|\sigma^{-m}(\mathcal{B}))$ denotes the expected projection (with respect to a initial probability P on the Bernoulii space) on the sigma-algebra $\sigma^{-m}(\mathcal{B})$, where $\mathcal{B}$ is the Borel sigma-algebra and $\Lambda^{[m]}$ is associated to the Jones index.

We refer the reader to [CL] for a Thermodynamic view of C^*-Algebras which include concepts like pressure, entropy, etc..

We believe it will be important in the analysis of certain C^* algebras associated to continuous time dynamical systems a characterization of KMS states by means of the above theorem. We point out, however, that we are able to show this theorem for a certain ρ_V just for a quite simple function V as above. In a forthcoming paper we will consider more general potentials V.

One could consider a continuous time version of the C^*-algebra considered in [EL1]. We just give an idea of what we are talking about. Given the above defined P for each $t > 0$, denote by $s_t : \mathcal{L}^2(P) \to \mathcal{L}^2(P)$, the Koopman operator, where for $\eta \in \mathcal{L}^2(P)$ we define $(s_t\eta)(x) = \eta(\theta_t(x))$.

Another important class of linear operators is $M_f : \mathcal{L}^2(P) \to \mathcal{L}^2(P)$, for a given fixed $f \in C(\Omega)$, and defined by $M_f(\eta)(x) = f(x)\eta(x)$, for any η in $\mathcal{L}^2(P)$. We assume that f is such defines an operator on $\mathcal{L}^2(P)$ (remember that Ω is not compact).

In this way we can consider a C^*-algebra generated by the above defined operators (for all different values of $t > 0$), then the concept of state, and finally given V and β we can ask about KMS states. There are several technical difficulties in the definition of the above C^*-algebra, etc... Anyway, at least formally, there is a need for finding ν which is a solution of an equation of the kind we describe here. We need this in order to be able to obtain a characterization of KMS states by means of an eigen-probability for the continuous time Ruelle operator. This setting will be the subject of a future work. This was the motivation for our result.

With the operators α and $\mathcal{L}$ we can rewrite the theorem above as

$$\rho_V(G_T^{-1}E_T(G_T\varphi)) = \rho_V(\varphi),$$

for all $\varphi \in \mathcal{L}^\infty$ and all $T > 0$, where, as usual, $\rho_V(\varphi) = \int \varphi d\rho_V$, $E_T = \alpha_T\mathcal{L}^T$ is in fact a projection on a subalgebra of $\mathcal{B}$, and $G_T : \Omega \to \mathbb{R}$ is given by

$$G_T(x) = \exp\,(\int_0^T V(x(s))ds).$$

For the map $V : \Omega \to \mathbb{R}$, which is constant in cylinders of the form $\{w_0 = i\}$, $i \in \{1, 2, ..., n\}$, we denote by V_i the corresponding value. We also denote by V the diagonal matrix with the i-diagonal element equal to V_i.

Now, consider $P_V^t = e^{t(L+V)}$. The classical Perron-Frobenius Theorem for such semigroup will be one of the main ingredients of our main proof.

As usual, we denote by $\mathcal{F}_s$ the sigma-algebra generated by X_s. We also denote by $\mathcal{F}_s^+$ the sigma-algebra generated $\sigma(\{X_u, s \leq u\})$. Note that a $\mathcal{F}_s^+$-measurable function $f(w)$ on Ω does depend of the value w_s.

We also denote by I_A the indicator function of a measurable set A in Ω.

2 A continuous time Ruelle Operator

We consider the disintegration of P given by the family of measures, indexed by the elements of Ω and $t > 0$ defined as follows: first, consider a sequence $0 = t_0 < t_1 < \cdots < t_{j-1} < t \leq t_j < \cdots < t_r$. Then for $w \in \Omega$ and $t > 0$ we have on cylinders:

$$\mu_t^w([X_0 = a_0, \ldots, X_{t_r} = a_r]) =$$

$$\begin{cases} \frac{1}{p_0^{w(t)}} P_{w(t)a_{j-1}}^{t-t_{j-1}} \cdots P_{a_2 a_1}^{t_2-t_1} P_{a_1 a_0}^{t} p_0^{a_0} & \text{if } a_j = w(t_j), \ldots, a_r = w(t_r) \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 2.1. *μ_t^w is the disintegration of P along the fibers $\Theta_t^{-1}(.)$.*

Proof: It is enough to show that for any integrable f

$$\int_\Omega f dP = \int_\Omega \int_{\Theta_t^{-1}(w)} f(x) d\mu_t^w(x) dP(w).$$

For doing that we can assume that f is in fact the indicator of the cylinder $[X_0 = a_0, \ldots, X_{t_r} = a_r]$; then the right hand side becomes

$$\int\int f d\mu_t^w(x) dP(w) = \int I_{[w(t_j)=a_j,\ldots,w(t_r)=a_r]} \frac{1}{p_0^{w(t)}} P_{w(t)a_{j-1}}^{t-t_{j-1}} \cdots P_{a_1 a_0}^{t_1} p_0^{a_0} dP(w) =$$

$$\sum_{a=1}^{n} \int I_{[w(t)=a, w(t_j)=a_j,\ldots,w(t_r)=a_r]} \frac{1}{p_0^{w(t)}} P_{w(t)a_{j-1}}^{t-t_{j-1}} \cdots P_{a_1 a_0}^{t_1} p_0^{a_0} dP(w) =$$

$$\sum_{a=1}^{n} P_{a_r a_{r-1}}^{t_r-t_{r-1}} \ldots P_{a_j a}^{t_j-t} p_0^a \frac{1}{p_0^a} P_{a a_{j-1}}^{t-t_{j-1}} \cdots P_{a_1 a_0}^{t_1} p_0^{a_0} = P_{a_r a_{r-1}}^{t_r-t_{r-1}} \cdots P_{a_1 a_0}^{t_1} p_0^{a_0} =$$

$$P([X_0 = a_0, \ldots, X_{t_r} = a_r]) = \int f dP.$$

In the second inequality we use the fact that P is stationary.

The proof for a general f follows from standard arguments. $\square$

Definition 2.2. *For t fixed we define the operator $\mathcal{L}^t : \mathcal{L}^\infty(\Omega, P) \to \mathcal{L}^\infty(\Omega, P)$ as follows:*

$$\mathcal{L}^t(\varphi)(x) = \int_{\bar{y} \in \Theta_t^{-1}(x)} \varphi(\bar{y}) d\mu_t^x(\bar{y}).$$

Remark 2.3. *The definition above can be rewritten as*

$$\mathcal{L}^t(\varphi)(x) = \int_{y\in D[0,t)} \varphi(yx) d\mu_t^x(yx),$$

where the symbol yx means the concatenation of the path y with the translation of x:

$$xy(s) = \begin{cases} y(s) & \text{if } s \in [0,t) \\ x(s-t) & \text{if } s \geq t, \end{cases}$$

and, $D[0,t)$ is the set of right-continuous functions from $[0,t)$ to S. This follows simply from the fact that, in this notation, $\Theta_t^{-1}(x) = \{yx \colon y \in D[0,t)\}$.

Note that the value $\lim_{s\to t} y(s)$ do not have to be necessarily equal to $x(0)$.

In order to understand better the definitions above we apply the operator to some simple functions. For example, we can see the effect of $\mathcal{L}^t$ on some indicator function of a given cylinder: consider the sequence $0 = t_0 < t_1 < .. < t_{j-1} < t \leq t_j < ... < t_r$ and then take $f = I_{\{X_0=a_0, X_{t_1}=a_1,...,X_{t_r}=a_r\}}$. Then, for a path $z \in \Omega$ such that $z_{t_j-t} = a_j, ..., z_{t_r-t} = a_r$ (the future condition) we have

$$\mathcal{L}^t(f)(z) = \frac{1}{p_0^{z_0}} P^{t-t_{j-1}}_{z_0 a_{j-1}} ... P^{t_2-t_1}_{a_2 a_1} P^{t_1}_{a_1 a_0} p_0^{a_0},$$

otherwise (i.e., if the path z does not satisfy the condition above) we get $\mathcal{L}^t(f)(z) = 0$.

Note that if $t_r < t$, then $\mathcal{L}^t(f)(z)$ depends only on z_0. For example, if $f = I_{\{X_0=i_0\}}$ then

$$\mathcal{L}^t(f)(z) = \int_{y\in D[0,t)} I_{\{X_0=i_0\}}(yx) d\mu_t^z(yx) = \mu_t^z([X_0 = i_0]) = \frac{1}{p_0^{z_0}} P^t_{z_0,i_0} p_0^{i_0}.$$

In the case $f = I_{\{X_0=i_0, X_t=j_0\}}$, then $\mathcal{L}^t(f)(z) = P^t_{z_0,i_0} \frac{p_0^{i_0}}{p_0^{z_0}}$, if $z_0 = j_0$, and $\mathcal{L}^t(f)(z) = 0$ otherwise.

We describe bellow some properties of $\mathcal{L}^t$.

Proposition 2.4. $\mathcal{L}^t(1) = 1$, *where* 1 *is the function that maps every point in* Ω *to* 1.

Proof: Indeed,

$$\mathcal{L}^t(1)(x) = \int_{y\in D[0,t)} 1(yx) d\mu_t^x(yx) = \int d\mu_t^x(yx) = \mu_t^x([X_t = x(0)]) =$$

$$\sum_{a=1}^n \mu_t^x([X_0 = a, X_t = x(0)]) = \frac{1}{p_0^{x(0)}} \sum_{a=1}^n P^t_{x(0)a} p_0^a = 1.$$

□

We can also define the dual of $\mathcal{L}^t$, denoted by $(\mathcal{L}^t)^*$, acting on the measures. Then we get:

Proposition 2.5. *For any positive t we have that $(\mathcal{L}^t)^*(P) = (P)$.*

Proof: For a fixed t we have that $(\mathcal{L}^t)^*(P) = (P)$, because for any f of the form $f = I_{\{X_0=a_0, X_{t_1}=a_1, ..., X_{t_r}=a_r\}}$, $0 = t_0 < t_1 < .. < t_{j-1} < t \leq t_j < ... < t_r$, we get

$$\int \mathcal{L}^t(f)(z)dP(z) = \sum_{b=1}^{n} \int_{\{X_0=b\}} \mathcal{L}^t(f)(z)dP(z) =$$

$$\sum_{b=1}^{n} \int I_{\{X_0=b, X_{t_j-t}=a_j, ..., X_{t_r-t}=a_r\}}(z)dP(z)\frac{1}{p_0^b} P_{ba_{j-1}}^{t-t_{j-1}} ... P_{a_2a_1}^{t_2-t_1} P_{a_1a_0}^{t_1} p_0^{a_0} =$$

$$\sum_{b=1}^{n} P(\{X_0 = b, X_{t_j-t} = a_j, X_{t_{j+1}-t} = a_{j+1}, ..., X_{t_r-t} = a_r\})\frac{1}{p_0^b} P_{ba_{j-1}}^{t-t_{j-1}} ... P_{a_2a_1}^{t_2-t_1} P_{a_1a_0}^{t_1} p_0^{a_0} =$$

$$\sum_{b=1}^{n} P_{a_ra_{r-1}}^{t_r-t_{r-1}} ... P_{a_{j+1}a_j}^{t_{j+1}-t_j} P_{a_jb}^{t_j-t} p_0^b \, \frac{1}{p_0^b} P_{ba_{j-1}}^{t-t_{j-1}} ... P_{a_2a_1}^{t_2-t_1} P_{a_1a_0}^{t_1} p_0^{a_0} =$$

$$\int f(w)dP(w).$$

$\square$

Proposition 2.6. *Given $t \in \mathbb{R}_+$, and the functions $\varphi, \psi \in \mathcal{L}^\infty(P)$, then*

$$\mathcal{L}^t(\varphi \times (\psi \circ \Theta_t))(z) = \psi(z) \times \mathcal{L}^t(\varphi)(z).$$

Proof:

$$\mathcal{L}^t(\varphi(\psi \circ \Theta_t))(x) = \int_{i\in D[0,t)} \varphi(ix)(\psi \circ \Theta_t)(ix)d\mu_t^x(i) =$$

$$\psi(x) \int_{i\in D[0,t)} \varphi(ix)d\mu_t^x(i) = (\psi\mathcal{L}^t(\varphi))(x) = \psi(x)\mathcal{L}^t(\varphi)(x),$$

since $\psi \circ \Theta_t(ix) = \psi(x)$, independently of i. $\square$

We just recall that the last proposition can be restated as

$$\mathcal{L}^t(\varphi\alpha_t(\psi)) = \psi\mathcal{L}^t(\varphi).$$

Then we get:

Proposition 2.7. *α_t is the dual of $\mathcal{L}^t$ on $\mathcal{L}^2(P)$.*

Proof: From last two propositions

$$\int \mathcal{L}^t(f)g\,dP = \int \mathcal{L}^t(f \times (g \circ \Theta_t))\,dP = \int f \times (g \circ \Theta_t)\,dP = \int f\alpha_t(g)dP,$$

as claimed. $\square$

We want to obtain conditional expectations in a more explicit form. For a given f, recall that the function $Z(w) = E(f|\mathcal{F}_t^+)$ is the Z (almost everywhere defined) $\mathcal{F}_t^+$-measurable function such that for any $\mathcal{F}_t^+$-measurable set B we have $\int_B Z(w)dP(w) = \int_B f(w)dP(w)$.

Proposition 2.8. *The conditional expectation is given by*

$$E(f|\mathcal{F}_t^+)(x) = \int f d\mu_t^x.$$

Proof: For t fixed, consider a $\mathcal{F}_t^+$-measurable set B. Then we have

$$\int_B \int f d\mu_t^w dP(w) = \int (I_B(w) \int f d\mu_t^w) dP(w) =$$

$$\int \int (f I_B) d\mu_t^w dP(w) = \int f(w) I_B(w) dP(w) = \int_B f dP.$$

$\square$

Now we can relate the conditional expectation with respect to the σ-algebras $\mathcal{F}_t^+$ with the operators $\mathcal{L}^t$ and α_t as follows:

Proposition 2.9. $[\mathcal{L}^t(f)](\Theta_t) = E(f|\mathcal{F}_t^+)$.

Proof: This follows from the fact that for any $B = \{X_{s_1} = b_1, X_{s_2} = b_2, ..., X_{s_u} = b_u\}$, with $t < s_1 < ... < s_u$, we have $I_B = I_A \circ \Theta_t$ for some measurable A and

$$\int_B \mathcal{L}^t(f)(\Theta_t(w))dP(w) = \int I_B(w)\mathcal{L}^t(f)(\Theta_t(w))dP(w) =$$

$$\int (I_A \circ \Theta_t)(w)\, \mathcal{L}^t(f)(\Theta_t(w))\, dP(w) = \int I_A(w)\mathcal{L}^t(f)(w)dP(w) =$$

$$\int I_A(\Theta_t(w))\, f(w)\, dP(w) = \int_B f(w)dP(w).$$

$\square$

3 The modified operator Ruelle Operator associated to V

We are interested in the perturbation by V (defined above) of the $\mathcal{L}^t$ operator.

Definition 3.1. *We define $G_t \colon \Omega \to \mathbb{R}$ as*

$$G_t(x) = \exp\,(\int_0^t V(x(s))ds)$$

Definition 3.2. *We define the G-weigthed transfer operator $\mathcal{L}_V^t : \mathcal{L}^\infty(\Omega, P) \to \mathcal{L}^\infty(\Omega, P)$ acting on measurable functions f (of the form $f = I_{\{X_0=a_0, X_{t_1}=a_1,\ldots,X_{t_r}=a_r\}}$) by*

$$\mathcal{L}_V^t(f)(w) := \mathcal{L}^t(G_t f) =$$

$$= \mathcal{L}^t(e^{\int_0^t (V\circ\Theta_s)(.)ds}\, f\,) = \sum_{b=1}^n \mathcal{L}^t(e^{\int_0^t (V\circ\Theta_s)(.)ds}\, I_{\{X_t=b\}}\, f\,)(w).$$

Note that $e^{\int_0^t (V\circ\Theta_s)(.)ds}\, I_{\{X_t=b\}}$ does not depend on information for time larger than t. In the case f is such that $t_r \leq t$ (in the above notation), then $\mathcal{L}_V^t(f)(w)$ depends only on $w(0)$.

The integration on s above is consider over the open interval $(0, t)$.

We will show next the existence of an eigenfunction and an eigen-measure for such operator $\mathcal{L}_V^t$. First we need the following:

Theorem (Perron-Frobenius for continuous time). *([S] page 111) Given L, p_0 and V as above, there exists*

a) a unique positive function $u_V : \Omega \to \mathbb{R}$, constant equal to the value u_V^i in each cylinder $X_0 = i$, $i \in \{1, 2, .., n\}$, (sometimes we will consider u_V as a vector in $\mathbb{R}^n$).

b) a unique probability vector μ_V in $\mathbb{R}^n$ (a probability over over the set $\{1, 2, .., n\}$ such that $\mu_V(\{i\}) > 0$, $\forall i$), that is,

$$\sum_{i=1}^n (u_V)_i\, (\mu_V)_i = 1,$$

c) a real positive value $\lambda(V)$, such that for any positive s

$$e^{-s\lambda(V)} u_V\, e^{s(L+V)} = u_V.$$

d) Moreover, for any $v = (v_1, ., v_n) \in \mathbb{R}^n$

$$\lim_{t\to\infty} e^{-t\lambda(V)}\, v\, e^{t(L+V)} = (\sum_{i=1}^n v_i(\mu_V)_i)\, u_V,$$

e) for any positive s

$$e^{-s\lambda(V)} e^{s(L+V)}\, \mu_V = \mu_V.$$

From property e) it follows that

$$(L+V)\,\mu_V = \lambda(V)\,\mu_V,$$

or

$$(L+V-\lambda(V)\,I)\,\mu_V \;=\; 0.$$

From c) it follows that

$$u_V\,(L+V) = \lambda(V)\,u_V,$$

or

$$u_v\,(L+V-\lambda(V)\,I) \;=\; 0.$$

We point out that e) means that for any positive t we have $(P_V^t)^*\,\mu_V = e^{\lambda(V)\,t}\mu_V$.

Note that when $V=0$, then $\lambda(V)=0$, $\mu_V = p^0$ and u_V is constant equal to 1.

Now we return to our setting: for each i_0 and t fixed one can consider the probability $\mu_{i_0}^t$ defined over the sigma-algebra $\mathcal{F}_t^- = \sigma(\{X_s | s \leq t\})$ with support on $\{X_0 = i_0\}$ such that for cylinder sets with $0 < t_1 < ... < t_r \leq t$

$$\mu_{i_0}^t(\{X_0 = i_0, X_{t_1} = a_1, ..., X_{t_{r-1}} = a_{r-1}, X_t = j_0\}) = P_{j_0 a_r}^{t-t_r} ... P_{a_2 a_1}^{t_2 - t_1} P_{a_1 i_0}^{t_1}.$$

The probability $\mu_{i_0}^t$ is not stationary.

We denote by $Q(j,i)_t$ the i,j entry of the matrix $e^{t(L+V)}$, that is $(e^{t(L+V)})_{j,i}$.

It is known ([K] page 52 or [S] Lemma 5.15) that

$$Q(j_0,i_0)_t = E_{\{X_0=i_0\}}\{e^{\int_0^t (V\circ\Theta_s)(w) ds}\,;\, X(t) = j_0\} =$$

$$\int I_{\{X_t=j_0\}}\; e^{\int_0^t (V\circ\Theta_s)(w) ds} d\mu_{i_0}^t(w).$$

For example,

$$\int I_{\{X_t=j_0\}} e^{\int_0^t (V\circ\Theta_s)(w) ds} dP = \sum_{i=1,2,..,n} Q(j_0,i)_t\, p_i^0.$$

In the particular case where V is constant equal 0, then $p^0 = \mu_V$ and $\lambda(V)=0$.

We denote by $f = f_V$, where $f(w) = f(w(0))$, the density of probability μ_V in S with respect to the probability p^0 in S.

Therefore, $\int f dp^0 = 1$.

Proposition 3.3. *$f_V(w) = \frac{\mu_V(w)}{p^0(w)} = \frac{(\mu_V)_{w(0)}}{(p^0)_{w(0)}}$, $f_V : \Omega \to \mathbb{R}$, is an eigenfunction for $\mathcal{L}_V^t$ with eigenvalue $e^{t\lambda(V)}$.*

Proof: Note that $\frac{\mu_V}{p^0} = \sum_{c=1}^n \frac{\mu_V(c)}{p^0(c)} I_{\{X_0=c\}}$.
For a given w, denote $w(0)$ by j_0, then conditioning

$$\mathcal{L}_V^t(\frac{\mu_V}{p^0})(w) = \sum_{c=1}^n \sum_{b=1}^n \mathcal{L}_V^t\,(\,\frac{\mu_V(c)}{p^0(c)}\, I_{\{X_0=c\}}\, I_{\{X_t=b\}}\,)\,(w).$$

Consider c fixed, then for $b = j_0$ we have

$$\mathcal{L}_V^t\,(\,I_{\{X_0=c\}}\, I_{\{X_t=b\}}\,)\,(w) = \frac{Q(j_0,c)_t\; p_c^0}{p_{j_0}^0},$$

and for $b \neq j_0$, we have $\mathcal{L}_V^t\,(\,I_{\{X_0=c\}}\, I_{\{X_t=b\}}\,)\,(w) = 0$.
Finally, for any $t > 0$

$$\mathcal{L}_V^t(\frac{\mu_V}{p^0})(w) = \sum_{c=1}^n \frac{\mu_V(c)}{p^0(c)}\, Q(j_0,c)_t\; \frac{p_c^0}{p_{j_0}^0} =$$

$$\sum_{c=1}^n \frac{\mu_V(c)}{p^0(c)}\,(e^{t(L+V)})_{j_0,c}\; \frac{p_c^0}{p_{j_0}^0} = e^{t\lambda(V)}\,\frac{(\mu_V)_{j_0}}{p_{j_0}^0} = e^{t\lambda(V)}(\frac{\mu_V}{p^0})(w),$$

because $e^{t(L+V)}(\,\mu_V) = e^{t\lambda(V)}(\mu_V)$.
Therefore, for any $t > 0$ the function $f_V = \frac{\mu_V}{p^0}$ (that depends only on $w(0)$) is an eigenfunction for the operator $\mathcal{L}_V^t$ associated to the eigenvector $e^{t\lambda(V)}$. $\square$

The above result shows that the eigenfunction for the Ruelle operator associated to the potential V is the Radon-Nykodin derivative for μ_V with respect to p^0;

Definition 3.4. *Consider for each t the operator acting on g given by*

$$\hat{\mathcal{L}}_V^t(g)(w) = [\,\frac{1}{f_V}\mathcal{L}^t(e^{\int_0^t (V-\lambda(V))\circ\Theta_s)(.)ds}\, g\, f_V)\,]\,(w)$$

From the above $\hat{\mathcal{L}}_V^t(1) = 1$, for all positive t.
We present some examples: note that by conditioning, if $g = I_{\{X_0=a_0, X_{t_1}=a_1, X_{t_2}=a_2, X_t=a_3\}}$, with $0 < t_1 < t_2 < t$, then

$$\hat{\mathcal{L}}_V^t(g)(w) = \frac{1}{p_{a_3}^0}\,\frac{p_{a_3}^0}{\mu_V(a_3)}\, e_{a_3a_2}^{(t-t_2)\,(L+V-\lambda\, I)}\, e_{a_2a_1}^{(t_2-t_1)\,(L+V-\lambda\, I)}\, e_{a_1a_0}^{t_1\,(L+V-\lambda\, I)}\,\frac{\mu_V(a_0)}{p_{a_0}^0}\, p_{a_0}^0 =$$

$$\frac{1}{\mu_V(a_3)}\, e_{a_3a_2}^{(t-t_2)\,(L+V-\lambda\, I)}\, e_{a_2a_1}^{(t_2-t_1)\,(L+V-\lambda\, I)}\, e_{a_1a_0}^{t_1\,(L+V-\lambda\, I)}\,\mu_V(a_0)\,,$$

for w such that $w_0 = a_3$, and $\hat{\mathcal{L}}_V^t(g)(w) = 0$ otherwise.

Moreover, for $g = I_{\{X_0=a_0, X_{t_1}=a_1, X_{t_2}=a_2\}}$, with $0 < t_1 < t < t_2$, then

$$\hat{\mathcal{L}}_V^t(g)(w) = \frac{1}{\mu_V(a)} e_{a\,a_1}^{(t-t_1)\,(L+V-\lambda\,I)}\, e_{a_1 a_0}^{t_1\,(L+V-\lambda\,I)}\, \mu_V(a_0),$$

for w such that $w_0 = a, w_{t_2-t} = a_2$, and $\hat{\mathcal{L}}_V^t(g)(w) = 0$ otherwise.

Consider now the dual operator $(\hat{\mathcal{L}}_V^t)^*$.

For t fixed consider the transformation in the set of measures μ on Ω given by $(\hat{\mathcal{L}}_V^t)^*(\mu) = \nu$.

Theorem 1. *There exists a fixed probability measure ν_V on $(\Omega, \mathcal{B})$ for such transformation $(\hat{\mathcal{L}}_V^t)^*$. The probability ν_V does not depend on t.*

Proof:

We have to show that there exists ν_V such that for all $t > 0$ and for all g we have

$$\int \hat{\mathcal{L}}_V^t(g)\, d\nu_V = \int g\, d\nu_V.$$

Remember that, $\hat{\mathcal{L}}_V^t(1) = 1$, therefore, if μ is a probability, then $(\hat{\mathcal{L}}_V^t)^*(\mu) = \nu$ is also a probability.

Denote by $\nu = \nu_V$ the probability obtained in the following way, for

$$g = I_{\{X_0=a_0, X_{t_1}=a_1, X_{t_2}=a_2, \ldots, X_{t_{r-1}}=a_{r-1}, X_r=a_r\}},$$

with $0 < t_1 < t_2 < \ldots < t_{s-1} < t_s < .. < t_r$, we define

$$\int g(w)\, d\nu(w) = \ e_{a_r a_{r-1}}^{(t_r-t_{r-1})\,(L+V-\lambda\,I)} \ldots e_{a_2 a_1}^{(t_2-t_1)\,(L+V-\lambda\,I)}\, e_{a_1 a_0}^{t_1\,(L+V-\lambda\,I)}\, \mu_V(a_0).$$

It is easy to see that this probability satisfies the Kolmogorov compatibility conditions.

In order to show that ν is a probability we have to use the fact that $\sum_{c\in S} \mu_V(c) = 1$

For example, $\int I_{\{X_0=c\}}\, d\nu = \mu_V(c)$. Moreover,

$$\int 1\, d\nu \ = \sum_c \sum_a \int I_{\{X_t=c, X_0=a\}}\, d\nu = \sum_c \sum_a e_{c\,a}^{t\,(L-V-\lambda\,I)}\, \mu_V(a) = \sum_c \mu_V(c) \ = \ 1.$$

Suppose t is such that $0 < t_1 < t_2 < \ldots < t_{s-1} < t \le t_s < .. < t_r$, then

$$z(w) = \hat{\mathcal{L}}_V^t(g)(w) = \frac{1}{\mu_V(a)} e_{a\,a_{s-1}}^{(t-t_{s-1})\,(L+V-\lambda\,I)} \ldots e_{a_2 a_1}^{(t_2-t_1)\,(L+V-\lambda\,I)}\, e_{a_1 a_0}^{t_1\,(L+V-\lambda\,I)}\, \mu_V(a_0),$$

for w such that $w(0) = a$, $w_{t_s-t} = a_s, w_{t_{s+1}-t} = a_{s+1}, \ldots, w_{t_r-t} = a_r$, and $\hat{\mathcal{L}}_V^t(g)(w) = 0$ otherwise. Note that $z(w) = \hat{\mathcal{L}}_V^t(g)(w)$ depends only on w_0, w_{t_s-t}, $w_{t_{s+1}-t}, \ldots, w_{t_r-t}$.

We have to show that for any g we have $\int g\, d\nu = \int \hat{\mathcal{L}}_V^t(g) d\nu$.

Now,

$$\int z(w)\,d\nu(w) = \int \sum_{c\in S} I_{\{X_0=c, X_{t_s-t}=a_s, X_{t_{s+1}-t}=a_{s+1},...,X_{t_r-t}=a_r\}}\, z(w)\;d\nu(w) =$$

$$\sum_{c\in S} \nu(\{X_0=c, X_{t_s-t}=a_s, X_{t_{s+1}-t}=a_{s+1},...,X_{t_r-t}=a_r\})$$

$$\frac{1}{\mu_V(c)} e^{(t-t_{s-1})\,(L+V-\lambda\,I)}_{c a_{s-1}} \,...\, e^{(t_2-t_1)\,(L+V-\lambda\,I)}_{a_2 a_1}\, e^{t_1\,(L+V-\lambda\,I)}_{a_1 a_0} \mu_V(a_0) =$$

$$\sum_{c\in S} e^{(t_r-t_{r-1})\,(L+V-\lambda\,I)}_{a_r a_{r-1}} \,...\, e^{(t_{s+1}-t_s)\,(L+V-\lambda\,I)}_{a_{s+1} a_s}\, e^{t_s-t\,(L+V-\lambda\,I)}_{a_s c}\, \mu_V(c)$$

$$\frac{1}{\mu_V(c)}\, e^{(t-t_{s-1})\,(L+V-\lambda\,I)}_{c a_{s-1}} \,...\, e^{(t_2-t_1)\,(L+V-\lambda\,I)}_{a_2 a_1}\, e^{t_1\,(L+V-\lambda\,I)}_{a_1 a_0} \mu_v(a_0) =$$

$$e^{(t_r-t_{r-1})\,(L+V-\lambda\,I)}_{a_r a_{r-1}} \,...\, e^{(t_{s+1}-t_s)\,(L+V-\lambda\,I)}_{a_{s+1} a_s}$$

$$(\sum_{c\in S} e^{(t_s-t)\,(L+V-\lambda\,I)}_{a_s c} e^{(t-t_{s-1})\,(L+V-\lambda\,I)}_{c a_{s-1}}) \,...\, e^{(t_2-t_1)\,(L+V-\lambda\,I)}_{a_2 a_1}\, e^{t_1\,(L+V-\lambda\,I)}_{a_1 a_0} \mu_V(a_0) =$$

$$,\, e^{(t_r-t_{r-1})\,(L+V-\lambda\,I)}_{a_r a_{r-1}} \,...\, e^{(t_{s+1}-t_s)\,(L+V-\lambda\,I)}_{a_{s+1} a_s}$$

$$e^{(t_s-t_{s-1})\,(L+V-\lambda\,I)}_{a_s a_{s-1}} \,...\, e^{(t_2-t_1)\,(L+V-\lambda\,I)}_{a_2 a_1}\, e^{t_1\,(L+V-\lambda\,I)}_{a_1 a_0}\, \mu(a_0) =$$

$$\int g\,d\nu.$$

The claim for the general g follows from the above result.
Therefore, $(\hat{\mathcal{L}}^t_V)^*(\nu_V) = \nu_V$.

□

Definition 3.5. *Consider the stationary probability $\rho_V = f_V\,\nu_V$ on Ω. We call it the equilibrium state for V.*

Definition 3.6. *We call the probability ν_V on Ω the Gibbs state for V.*

Proposition 3.7. *For any integrable $f, g \in \mathcal{L}^\infty(P)$ and any positive t*

$$\int \hat{\mathcal{L}}^t_V(f)\,g\,d\nu_V = \int \hat{\mathcal{L}}^t_V(f(g\circ\theta_t))\,d\nu_V = \int f(g\circ\theta_t)\,d\nu_V.$$

Now we can prove our main result.

Theorem A. *For any integrable $g \in \mathcal{L}^\infty(P)$ and any positive t*

$$\int e^{-\int_0^t (V\circ\Theta_s)(.)ds}\;[\,(\frac{1}{f_V}\,\mathcal{L}^t\,(e^{\int_0^t (V\circ\Theta_s)(.)ds} g\,f_V)\,)\,\circ\theta_t\,]\,d\nu_V = \int g d\nu_V.$$

Proof:
Note first that if w is such that $w(0)=c$, then

$$\mathcal{L}^t(f_V(w)) = \sum_j \mathcal{L}^t\,(I_{\{X_0=j\}}\,)\, f_V(j) = \sum_j \mathcal{L}^t\,(I_{\{X_0=j\}}\,)\frac{\mu_V(j)}{p_j^0} =$$

$$\sum_j \frac{1}{p_c^0}\,P_{c\,j}\,p_j^0\,\frac{\mu_V(j)}{p_j^0} = \sum_j \frac{1}{p_c^0}\,P_{c\,j}\,\mu_V(j) = \frac{\mu_V(c)}{p_c^0} = f_V(w).$$

Therefore, for all w

$$\frac{1}{f_V}\,\mathcal{L}^t(f_V(w)) = 1.$$

Finally,

$$\int e^{-\int_0^t (V\circ\Theta_s)(.)ds}\;[\,(\frac{1}{f_V}\,\mathcal{L}^t\,(e^{\int_0^t (V\circ\Theta_s)(.)ds} g\, f_V)\,)\;\circ\theta_t\,]d\nu_V =$$

$$\int e^{-\int_0^t (V\circ\Theta_s-\lambda)(.)ds}\;[\,(\frac{1}{f_V}\,\mathcal{L}^t\,(e^{\int_0^t (V\circ\Theta_s-\lambda)(.)ds} g\, f_V)\,)\;\circ\theta_t\,]d\nu_V =$$

$$\int [\hat{\mathcal{L}}_V^t(e^{-\int_0^t (V\circ\Theta_s-\lambda)(.)ds})]\,[\,\frac{1}{f_V}\mathcal{L}^t\,(e^{\int_0^t (V\circ\Theta_s-\lambda)(.)ds} g\, f_V)\,]\,d\nu_V =$$

$$\int [\,\frac{1}{f_V}\,\mathcal{L}^t\,(e^{-\int_0^t (V\circ\Theta_s)(.)ds}\,e^{\int_0^t (V\circ\Theta_s)(.)ds}\;f_V)\,]\,[\frac{1}{f_V}\,\mathcal{L}^t\,(e^{\int_0^t (V\circ\Theta_s-\lambda)(.)ds}\;f_V\,g)\,]\,d\nu_V =$$

$$\int [\,\frac{1}{f_V}\,\mathcal{L}^t(f_V)\,]\;[\,\frac{1}{f_V}\,\mathcal{L}^t\,(e^{\int_0^t (V\circ\Theta_s-\lambda)(.)ds} f_V\,g)\,]\,d\nu_V \;=\; \int \frac{1}{f_V}\;\mathcal{L}^t\,(e^{\int_0^t (V\circ\Theta_s-\lambda)(.)ds}\;f_V\,g)\,d\nu_V \;=$$

$$\int \hat{\mathcal{L}}_V^t(g)\,d\nu_V \;=\; \int g d\nu_V.$$

$\square$

Bibliography

[B] P. Billingsley, Convergence of probability measures, John Wiley, 1968

[CL] Gilles G. de Castro and Artur O. Lopes, KMS States, Entropy and a Variational Principle for Pressure, preprint 2007

[BW] R. Bhattacharya and E. Waymire, Stochastic Processes with applications, Wiley (1990)

[EK] S. Ethier and T. Kurtz, Markov Processes, John Wiley, 1986

[EL1] R. Exel and A. O. Lopes, "C^*-Algebras, approximately proper equivalence relations, and Thermodynamic Formalism", *Erg. Theo. and Dyn. Systems.* **24** (2004) 1051-1082

[EL2] R. Exel and A. O. Lopes, "C^*-Algebras and Thermodynamic Formalism", to appear in *São Paulo Journal of Mathematics.*

[K] M. Kac, Integration in Function spaces and some of its applications, Acad Naz dei Lincei Scuola Superiore Normale Superiore, Piza, Italy (1980).

[N] J. B. Norris, Markov Chains, Cambridge Press

[P] K. Parthasarathy, Probability measures on metric spaces, Academic Press,

[S] D. W. Strook, An introduction to Large Deviations, Springer , 1984